Amendment to the Specification:

Please amend the specification as follows:

Please replace the paragraph [0001] on page 1, with the following amended paragraph:

[0001] The present application is a continuation-in-part application that claims priority from U.S. Application Serial No. [[09/412,894]] 09/412,184, filed October 4, 1999, issued as U.S. Patent No. 6,268,188 [[allowed]], and U.S. Application Serial No. 09/389,537, now pending, filed September 2, 1999; both of which are continuations of U.S. Application Serial No. 08/646,590, filed May 8, 1996, now U.S. Patent 5,962,283; which is a continuation-in-part of U.S. Application Serial No. 08/599,171, filed February 9, 1996, now U.S. Patent 5,814,473; and U.S. Application no. 09/481,733, filed January 11, 2000, now pending; which is a continuation of U.S. Application No. 09/069,226, filed April 27, 1998, now U.S. Patent 6,013,509; which is a continuation of U.S. Application Serial No. 08/599,171, filed February 9, 1996, now U.S. Patent 5,814,473, the contents of which are hereby incorporated by reference in their entirety.

Please replace the paragraph [0036] on page 6, with the following amended paragraph:

[0036] Figure 9 is an illustration of the full-length DNA (SEQ ID NO: 21) and corresponding deduced amino acid sequence (SEQ ID NO: 29) of Ammonifex [[Aquifex]] degensii aspartate aminotransferase.

Please replace the paragraph [0063] on page 13, with the following amended paragraph:

[0063] The term "variant" refers to polynucleotides or polypeptides of the invention modified at one or more base pairs, codons, introns, exons, or amino acid residues (respectively) yet still retain the biological activity of a transaminase or aminotransferase of the invention. Variants can be produced by any number of means included methods such as, for example, error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble

mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly,

[[GSSM]] gene site saturated mutagenesis (GSSM™) and any combination thereof.

      Please replace the paragraph [0139] on page 31, with the following amended paragraph:

[0139] The invention also provides for the use of proprietary codon primers (containing a degenerate N,N,N sequence) to introduce point mutations into a polynucleotide, so as to generate a set of progeny polypeptides in which a full range of single amino acid substitutions is represented at each amino acid position (gene site saturated mutagenesis [[(GSSM)]] (GSSM)™). The oligos used are comprised contiguously of a first homologous sequence, a degenerate N,N,N sequence, and preferably but not necessarily a second homologous sequence. The downstream progeny translational products from the use of such oligos include all possible amino acid changes at each amino acid site along the polypeptide, because the degeneracy of the N,N,N sequence includes codons for all 20 amino acids.

      Please replace the paragraph [0220] on page 31, with the following amended paragraph:

[0220] The polypeptides of Group B amino acid sequences, and sequences substantially identical thereto or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof may be used in a variety of applications. For example, the polypeptides or fragments thereof may be used to catalyze biochemical reactions. In accordance with one aspect of the invention, there is provided a process for utilizing the polypeptides of Group B amino acid sequences, and sequences substantially identical thereto or polynucleotides encoding such polypeptides for transamination reactions hydrolyzing glycosidic linkages. In such procedures, a substance containing an amino acid a glycosidic linkage (e.g., a starch) is contacted with one of the polypeptides of Group B amino acid sequences, or sequences substantially identical thereto under conditions which facilitate the transamination reaction hydrolysis of the glycosidic linkage.

Applicant : Patrick V. Warren et al.   Attorney's Docket No.: 09010-017006 / DIVER 1240-7
Serial No. : 09/905,173
Filed   : July 12, 2001
Page   : 4 of 20

Please replace the paragraph [0242] on page 60 to 61, with the following amended

paragraph:

[0242] A "comparison window", as used herein, includes reference to a segment

of any one of the number of contiguous positions selected from the group consisting of from 20

to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence

may be compared to a reference sequence of the same number of contiguous positions after the

two sequences are optimally aligned. Methods of alignment of sequence for comparison are

well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g.,

by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482, 1981, by the

homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol 48:443, 1970, by the

search for similarity method of person & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444, 1988, by

computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in

the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr.,

Madison, WI), or by manual alignment and visual inspection. Other algorithms for determining

homology or identity include, for example, in addition to a BLAST program (Basic Local

Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS

(Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment),

ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN

(Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher),

FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS,

LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm,

FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER,

FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL,

GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence

Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction &

Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-

Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) and

WHAT-IF. Such alignment programs can also be used to screen genome databases to identify

polynucleotide sequences having substantially identical sequences. A number of genome

databases are available, for example, a substantial portion of the human genome is available as

part of the Human Genome Sequencing Project ~~(J. Roach,~~

~~http://weber.u.Washington.edu/~roach/human_genome_progress 2.html) (Gibbs, 1995)~~. At

least twenty-one other genomes have already been sequenced, including, for example, M.

genitalium (Fraser et al., 1995), M. jannaschii (Bult et al., 1996), H. influenzae (Fleischmann et

al., 1995), E. coli (Blattner et al., 1997), and yeast (S. cerevisiae) (Mewes et al., 1997), and D.

melanogaster (Adams et al., 2000). Significant progress has also been made in sequencing the

genomes of model organism, such as mouse, C. elegans, and Arabadopsis sp. Several databases

containing genomic information annotated with some functional information are maintained by

different organization, and are accessible via the internet~~, for example, http://wwwtigr.org/tdb;~~

~~http://www.genetics.wisc.edu; http://genome-www.stanford.edu/~ball; http://hiv-web.lanl.gov;~~

~~http://www.ncbi.nlm.nih.gov; http://www.ebi.ac.uk; http://Pasteur.fr/other/biology; and http://~~

~~www.genome.wi.mit.edu.~~

Please replace the paragraph [0243] on page 61 to 62, with the following amended

paragraph:

[0243] One example of a useful algorithm is BLAST and BLAST 2.0 algorithms,

which are described in Altschul et al., Nuc. Acids Res. 25:3389-3402, 1977, and Altschul et al.,

J. Mol. Biol. 215:403-410, 1990, respectively. Software for performing BLAST analyses is

publicly available through the National Center for Biotechnology Information

~~(http://www.ncbi.nlm.nih.gov/)~~. This algorithm involves first identifying high scoring sequence

pairs (HSPs) by identifying short words of length W in the query sequence, which either match

or satisfy some positive-valued threshold score T when aligned with a word of the same length in

a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al.,

supra). These initial neighborhood word hits act as seeds for initiating searches to find longer

HSPs containing them. The word hits are extended in both directions along each sequence for as

far as the cumulative alignment score can be increased. Cumulative scores are calculated using,

for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always

>0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score.

Extension of the word hits in each direction are halted when: the cumulative alignment score

falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero

or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectations (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) alignments (B) of 50, expectation (E) of 10, M=5, N= -4, and a comparison of both strands.

Please replace the paragraph [0251] on page 62 to 63, with the following amended paragraph:

[0251] The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (i.e., aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet et al., Science 256:1443-1445, 1992; Henikoff and Henikoff, Proteins 17:49-61, 1993). Less preferably, the PAM or PAM250 matrices may also be used (see, e.g., Schwartz and Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation). BLAST programs are accessible through the U.S. National Library of Medicine, ~~e.g., at www.ncbi.nlm.nih.gov~~.

Please replace the paragraph [0268] on page 67 to 68, with the following amended paragraph:

[0268] Figure 4 is a flow diagram illustrating one embodiment of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a

list of each feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An example of such a database is produced by the University of Wisconsin Genetics Computer Group (~~www.gcg.com~~). Alternatively, the features may be structural polypeptide motifs such as alpha helices, beta sheets, or functional polypeptide motifs such as enzymatic active sites, helix-turn-helix motifs or other motifs known to those skilled in the art.

       Please replace the paragraph [0291] on page 75, with the following amended paragraph:

Ammonifex degensii [[hp]] Histadine-phosphate aminotransferase

[0291] 5'-

CCGAGAATTCATTAAAGAGGAGAAATTAACTATGGCAGTCAAAGTGCGGCCT
        (SEQ ID NO:33).

3'-CGGAGGATCCTTATCCAAAGCTTCCAGGAAG (SEQ ID NO:34).